How are Ego-centric Networks of Researchers Coupled?

Yaxue Ma [0000-0002-3680-3222], Jin Mao* [0000-001-9572-6709] and Gang Li [0000-0001-5573-6400]

Center for Studies of Information Resources, Wuhan University, Wuhan, China
{myx_vicky,maojin,ligang}@whu.edu.cn

Abstract. Scientific knowledge creation can be viewed as social-economic activities, which inspires us to explore researchers' interpersonal capital and its impact on scientific performance. In this study, we investigate on multiple types of interpersonal relationships between researchers, including co-authorship, author citation, and social relation, which are considered as interpersonal capital of researchers. Thus, three types of ego-centric networks (ECNs) are constructed by using the data from Twitter and Web of Science. The composition of social networks and the coupling relationships between ECNs in terms of the same researchers are analyzed. The preliminary results on the field of Cheminformatics show that most researchers tend to interact with research related accounts in social networks. The coupling degree between co-authorship networks and author citation networks is significantly higher than that between co-authorship networks and friend networks. Researchers are more likely to collaborate with the researchers who have close scholarly communication with them than the friends from social networks. This study contributes to the understanding of interpersonal relationship in scientific community. Future research will focus on the impact of interpersonal capital on scientific performance.

Keywords: Ego-centric Network, Scientific Production, Interpersonal Capital.

1 Introduction

Interpersonal relations are regarded as the main part of social capital in economic activities [1]. Interpersonal capital refers to the accumulation of interpersonal relations which contributes to future production or exchange [2]. In many social activities, interpersonal capital has been studied as important labor elements. Scientific knowledge creation in the academic environment is essentially a collective and social activity [3]. Especially in modern science, a growing number of researchers are communicating and collaborating with others to exchange new ideas or share technical resources [4]. It has been proven that scientific collaboration can foster the efficiency of the scientific production process and shorten the time of obtaining research results attribute to the rational division of scientific labor [5]. These are the situations that interpersonal capital takes effect in academic activities.

In social network analysis (SNA), the interpersonal capital of one actor is often demonstrated by its connections with other actors. To analyze the interpersonal capital, the ego-centric networks (ECNs) can be constructed by modeling the relationships between the target actor (ego) and other connected actors (alters). For a researcher,

ego-centric networks based on the relationships between him/her (ego) and other researchers (alters) can reflect interpersonal capital [6].

In the real academic environment, the relationships between researchers are diverse and multi-faceted. Depending on the type of relationship, different ECNs of researchers can be constructed. The two most common ECNs that are based on scholarly communication are co-authorship network and author citation network. Co-authorship networks rely on the co-authorship relations between researchers, while author citation networks are based on the citations between researchers' papers. Out of formal scholarly communication through publications, researchers may also have some other social connections, such as friendship, colleagueship, etc. Social network sites have been widely used for informal scholarly communication [7]. In this study, we are interested in researchers' interactions in social network sites (e.g., Twitter), and constructed social ECNs. In sum, three types of ECNs are investigated in this study including co-authorship network, author citation network, and social network. Intuitively, we raise two research questions:

• What are the characteristics of researchers' ego-centric social networks? Are a large number of alters engaged in research related occupations, in other words, is the composition of ego-centric social networks highly homogenized?

• Are there coupling relations between researchers' three ECNs in terms of the same researchers in these networks? It is not rare that the users in social network sites are the co-authors of papers. Similarly, the researchers citing or cited by a researcher could become collaborators. We will investigate how are the three ECNs coupled.

In this poster, we report our on-going project on analyzing researchers' ECNs in the field of Cheminformatics, which is the first attempt of our study. The three ECNs could influence the process of scientific production and potentially benefit researchers' scientific performance. The conception model could be seen as Fig. 1. A few recent studies have explored the relationship between the attributes of co-author network and scientific performance from the perspective of network structure [8-10]. Mccarty introduced the h-index of co-authors as a factor that may affect scientific performance [11], the focus of their research is still on network structure. However, few studies focus on the relationship between the attributes of other researchers (i.e., alters in ECN) and the researcher's scientific performance. That will be explored in the near future.

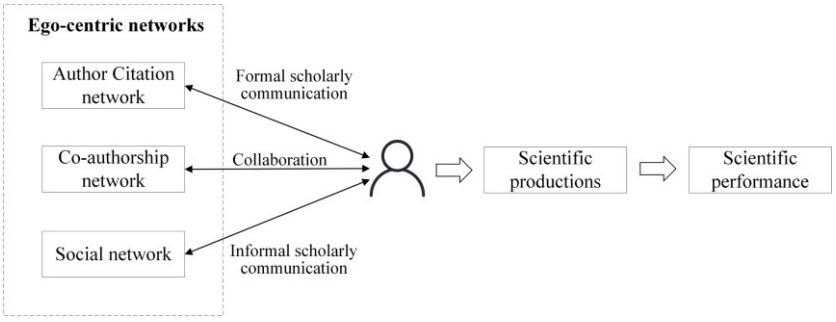

Fig. 1. The relationship between researchers' ego-centric networks and scientific performance.

2 Data and Methodology

2.1 Data Collection

Twitter, as a world-wide popular social network site, blurs the boundaries between profession and individual [12]. Many researchers regard their profiles as a way to boost their professional presence online, and most of them tend to post content about research work [13]. Therefore, Twitter is used in this study to obtain social relationships of researchers. Researchers in the field of Cheminformatics were chosen to be analyzed in this study in that many researchers in the field were active Twitter users [14].

We constructed co-authorship networks, author citation networks, and social networks of the researchers in this field as the following steps:

1) To get candidate researchers, *Cheminformatics* was used as the topic keyword to retrieve related literature from the Web of Science (WoS) Core Collection. In total, metadata of 786 articles and reviews was downloaded and parsed to obtain author information. Then, 489 authors with at least two publications were selected as candidates.

2) The names of candidate researchers were used as queries to search Twitter for possible user accounts. The affiliations, personal homepages, and avatars of the candidates were collected and referenced to manually determine the final accounts of the candidates. This process successfully matched 40 researchers as egos. Twitter API was applied to get the followers and followings of the 40 researchers, which compose the alters of the egos in social networks. Then, the friendship relations were found for the egos and alters who follow each other. And, the account names and personal descriptions of both egos and alters were obtained through the API.

3) To complete the publications of the 40 researchers, their names were used to retrieve articles and reviews from WoS. Due to the issue of author name ambiguation, we used the afflictions and ORCIDs of the researchers to filter out irrelevant publications. Co-authorship networks were formed based on the publications. The papers that cited the publications of the researchers were also fetched from WoS. Then, the references of the egos' publications and the papers citing the egos were used to construct author citation networks, where the links indicate that the authors cite each other.

2.2 Methodology

Identify Research Related Alters in Researchers' Ego-centric Social Network. We investigate the composition of researchers' social networks by identifying research-related alters. A list of terms was compiled by including terms such as "professor", "university", etc. The term list was used to obtain potential researchers by matching the personal descriptions of users in social networks. We built a small test set with 500 alters to evaluate this classification method. The accuracy is 92.60%, which evidences that our method is feasible. Some accounts of research institutions and teams were also found by this method, which are also treated as research related

alters. For users without personal descriptions, they were judged by matching their names with the authors in co-authorship networks and author citation networks.

We define the degree of homogeneity for social networks as the proportion that researchers occupy. The score of homogeneity is calculated as E-I index [15]:

$$Homogeneity = \frac{I\text{-}E}{I+E} \quad (1)$$

where, I is the number of research-related alters, and E is that of non-related alters. The higher the score of homogeneity is, the more alters are engaged in scientific research in the network.

The Coupling Degree between Co-authorship Networks, Author Citation Networks, and Social Networks. It is evidenced that friends tend to work in the same field and have similar ideas [16]. Further, we would like to explore the coupling degree between the three types of ego-centric networks to investigate how the three ECNs overlap with each other. Name string matching was applied to identify coupled alters (i.e. the same researchers) between the three ECNs. This method is reasonable in that 98.59% of individual users with personal descriptions in Twitter use real names according to our statistical report. The coupling degree of two networks is measured as:

$$coupling\ degree = \frac{|U_{net_1} \cap U_{net_2}|}{|U_{net_1}|} \quad (2)$$

where, U_{net_n} denotes the node set of the network net_n that could be one of the three ECNs, $|U_{net_n}|$ is the size of the node set net_n, while $|U_{net_1} \cap U_{net_2}|$ is the number of coupled nodes in the two networks.

3 Results

3.1 Results of ECNs

The basic information of the 40 researchers in the field of Cheminformatics is shown in Table 1. The alters that cite the egos and are cited by the egos are over 10 times than the alters of other relations. It has been shown that reciprocal citations indicate some close academic connections between authors [17]. Thus, we define our author citation networks based on reciprocal citations in this study.

Table 1. Basic information of ECNs for the 40 researchers in Cheminformatics.

Egos	Avg. co-authors	Avg. authors that are cited by the egos	Avg. authors that cited the egos	Avg. authors that cited each other with egos	Avg. users that the egos follows	Avg. users that follow the egos	Avg. of friends
40	274.03	9322.73	6326.70	908.05	568.55	726.48	206.63

3.2 The compositions of social ECNs

The homogeneity of the follower networks, following networks, friend networks where the users follow each other, and comprehensive networks with all types of alters, are presented for the 40 researchers in Fig. 2. It can be found that the number of alters in social networks with homogeneity greater than 0 is more than those with homogeneity smaller than 0. The homogeneity scores of friend networks are significantly higher than those of the other networks. It means that most researchers in Cheminformatics tend to interact with research related users in social network sites, especially follow each other. It should be noted that some researchers have diverse social relations, whose homogeneity scores are smaller than 0.

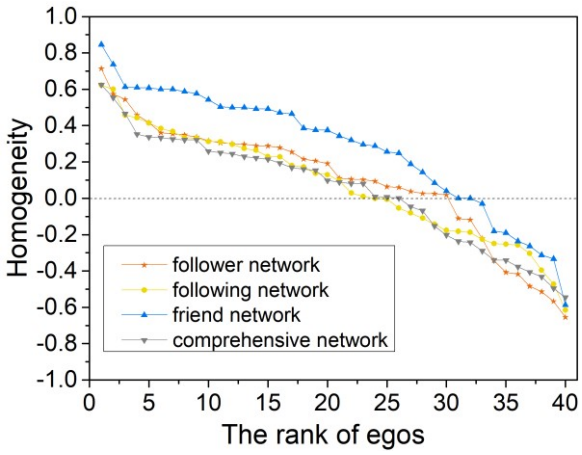

Fig. 2. The homogeneity scores of social ECNs.

Table 2. The distribution of the coupling degrees.

Coupling degree range	Co-authorship network v.s. Friend network		Co-authorship network v.s. Author citation network	
	# of Egos	Prop.	# of Egos	Prop.
0	5	12.50%	0	-
（0,0.2)	33	82.50%	16	40.00%
[0.2,0.4)	1	2.50%	19	47.50%
[0.4,0.6)	1	2.50%	3	7.50%
[0.6,0.8)	0	-	1	2.50%
[0.8,1)	0	-	0	-
1	0	-	1	2.50%

3.3 The coupling degree of ECNs

Table 2 reports the coupling degrees between co-authorship networks and author citation networks as well as friend networks. For co-author networks and friend networks, the coupling degree of most egos (95%) is lower than 0.2. It reflects that most friends in social networks will not collaborate. By contrast, the coupling degree between co-authorship networks and author citation networks are relatively higher. In total, 57.5% of the egos have a coupling degree from 0.2 to 0.8, even 2.5% of the egos achieve a degree of 1.0. The result means that the researchers are more likely to collaborate with those who have close scholarly communication with them than friends from social networks.

4 Conclusion

Taking the field of Cheminformatics as an example, this study presents preliminary results on researchers' ego-centric networks, including co-authorship networks, author citation networks, and social networks. As scholars, many researchers will follow or be followed by other researchers, which could due to the characteristics of academic occupations. The coupling results between co-authorship networks and author citation networks as well as friend networks indicate that formal scholarly communication through publications could bring a higher chance of publication than informal communication through social network interaction. These results should be examined in more research fields. As a preliminary attempt to explore researchers' interpersonal capital, we hope to provide some ideas and experiences for future studies. The next

step of our research is to further explore the impact of ego-centric networks on researchers' scientific performance by analyzing more disciplines.

Acknowledgments. This study was funded by the NSFC (Nos. 71804135).

References

1. Granovetter M S.: The Strength of Weak Ties. American Journal of Sociology 78(6), 1360-1380 (1973).
2. Wang, X.: Economists in China should pay more attention to the issue of interpersonal relations. Research on Financial and Economic Issues (6), 3-11 (2000).
3. Cannella, J. R., Albert, A., Mcfadyen, M. A.: Changing the exchange: The dynamics of knowledge worker ego networks. Journal of Management 42(4), 1005-1029 (2016).
4. Katz, J. S., Martin, B. R.: What is research collaboration?. Research Policy 26(1), 1-18 (1997).
5. Barrios, C., Flores, E., Martínez, M. Á., et al.: Is there convergence in international research collaboration? An exploration at the country level in the basic and applied science fields. Scientometrics, 1-29 (2019).
6. Chung, K. K., Hossain, L., Davis, J.: Exploring sociocentric and egocentric approaches for social network analysis. In: Proceedings of the 2nd international conference on knowledge management in Asia Pacific, pp. 1-8. (2005).
7. Kjellberg, S., Haider, J., Sundin, O.: Researchers' use of social network sites: A scoping review. Library & Information Science Research 38(3), 224-234 (2016).
8. Abbasi, A., Chung, K. S. K., Hossain, L.: Egocentric analysis of co-author network structure, position and performance. Information Processing & Management 48(4), 671-679 (2012).
9. Li, E. Y., Liao, C. H., Yen, H. R.: Co-author networks and research impact: A social capital perspective. Research Policy 42(9), 1515-1530 (2013).
10. Wang, J.: Knowledge creation in collaboration networks: Effects of tie configuration. Research Policy 45(1), 68-80 (2016).
11. Mccarty, C., Jawitz, J. W., Hopkins, A., Goldman, A.: Predicting author h-index using characteristics of the co-author network. Scientometrics 96(2), 467-483 (2013).
12. Bornmann, L.: Usefulness of altmetrics for measuring the broader impact of research. Aslib Journal of Information Management 67(3), 305-319 (2015).
13. Van Noorden, R.: Online collaboration: Scientists and the social network. Nature news 512(7513), 126 (2014).
14. Holmberg, K., Thelwall, M.: Disciplinary differences in Twitter scholarly communication. Scientometrics 101(2), 1027-1042 (2014).
15. Krackhardt, D., Stern, R. N.: Informal networks and organizational crises: An experimental simulation. Social psychology quarterly 51(2), 123-140 (1988).
16. White, H. D., Wellman, B., Nazer, N.: Does citation reflect social structure? : Longitudinal evidence from the "Globenet" interdisciplinary research group. Journal of the American Society for information Science and Technology 55(2), 111-126 (2004).
17. Daud, A., Ahmed, W., Amjad, T., et al.: Who will cite you back? reciprocal link prediction in citation networks. Library Hi Tech 35(3), 00-00 (2017).